



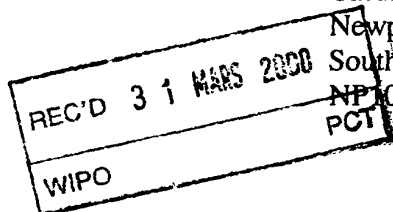
The
Patent
Office

PCT/US 99/30503



INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ
PCT

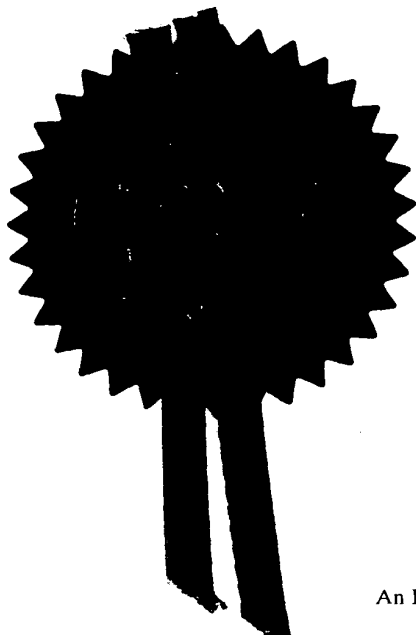


I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



Signed *Andrew Gacey*
Dated 5 January 2000

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)



DEC98 E414032-2 002882
9817700 0.00 - 9828377.3

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1.	Your reference	SCB/51935/000		
2.	Patent application number (The Patent Office will fill in this part)	9828377.3		
3.	Full name, address and postcode of the or of each applicant (<i>underline all surnames</i>)	JANSSEN PHARMACEUTICA N.V. TURNHOUTSEWEG 30 B-2340 BEERSE BELGIUM		
	Patents ADP number (<i>if you know it</i>)	531939001		
	If the applicant is a corporate body, give the country/state of its incorporation	BELGIUM		
4.	Title of the invention	VASCULAR ENDOTHELIAL GROWTH FACTOR-E		
5.	Name of your agent (<i>if you have one</i>)	BOULT WADE TENNANT		
	"Address for service" in the United Kingdom to which all correspondence should be sent (<i>including the postcode</i>)	27 FURNIVAL STREET LONDON EC4A 1PQ		
	Patents ADP number (<i>if you know it</i>)	42001		
6.	If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (<i>if you know it</i>) the or each application number	Country	Priority application number (if you know it)	Date of filing (day/month/year)
7.	If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	Number of earlier application	Date of filing (day / month / year)	
8.	Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if: a) any applicant named in part 3 is not an inventor, or b) there is an inventor who is not named as an applicant, or c) any named applicant is a corporate body. See note (d))	YES		

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form 0

Description 14

Claim(s) 4

Abstract 0

Drawing(s) 12

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77)

Request for preliminary examination and search (Patents Form 9/77)

Request for substantive examination (Patents Form 10/77)

Any other documents (Please specify)

11.

I/We request the grant of a patent on the basis of this application.

Signature

Date

Colm D. Murphy

22 December 1998

12. Name and daytime telephone number of person to contact in the United Kingdom

COLM D. MURPHY
0171 404 5921

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

- If you need help to fill in this form or you have any questions, please contact the Patent Office on 01645 500505.
- Write your answers in capital letters using black ink or you may type them.
- If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- Once you have filled in the form you must remember to sign and date it.
- For details of the fee and ways to pay please contact the Patent Office.

VASCULAR ENDOTHELIAL GROWTH FACTOR-E

The present invention is concerned with a novel
vascular endothelial growth factor (VEGF) herein
5 designated "VEGF-E", and characterisation of the
nucleic acid and amino acid sequences of VEGF-E.

Angiogenesis involves formation and proliferation of
new blood vessels, and is an essential physiological
10 process for normal growth and development of tissues
in, for example, embryonic development, tissue
regeneration and organ and tissue repair.
Angiogenesis also features in the growth of human
cancers which require continuous stimulation of blood
15 vessel growth. Abnormal angiogenesis is associated
with other diseases such as rheumatoid arthritis and
psoriasis.

Capillary vessels consist of endothelial cells which
20 carry the genetic information necessary to proliferate
to form capillary networks. Angiogenic molecules
which can initiate this process have previously been
characterised. A highly selective mitogen for
vascular endothelial cells is vascular endothelial
25 growth factor (VEGF) (Ferrara et al., "Vascular
Endothelial Growth Factor: Basic Biology and Clinical
Implications". Regulation of angiogenesis, by I.D.
Goldberg and E.M. Rosen 1997 Birkhauser Verlag
Basle/Switzerland). VEGF is a potent vasoactive
30 protein which is comprised of a glycosylated cationic
46-49 kd dimer having two 24 kd subunits. It is
inactivated by sulfhydryl reducing agents and is
resistant to acidic pH and to heating and binds to
immobilised heparin.

VEGF has four different forms of 121, 165, 189 and 206 amino acids due to alternative splicing. VEGF121 and VEGF165 are soluble and are capable of promoting angiogenesis, whereas VEGF189 and VEGF206 are bound to heparin containing proteoglycans in the cell surface. The temporal and spatial expression of VEGF has been correlated with physiological proliferation of the blood vessels (Gajdusek, C.M., and Carbon, S.J., *Cell Physiol.*, 139:570-579, (1989)); McNeil, P.L., Muthukrishnan, L., Warder, E., D'Amore, P.A., *J. Cell. Biol.*, 109:811-822, (1989)). Its high affinity binding sites are localized only on endothelial cells in tissue sections (Jakeman, L.B., et al., *Clin. Invest.* 89:244-253, (1989)). The growth factor can be isolated from pituitary cells and several tumor cell lines, and has been implicated in some human gliomas (Plate, K.H. *Nature* 359:845-848, (1992)). The inhibition of VEGF function by anti-VEGF monoclonal antibodies was shown to inhibit tumor growth in immune-deficient mice (Kim, K.J., *Nature* 362:841-844, (1993)).

The present inventors have now identified a further vascular endothelial growth factor, designated herein as "VEGF-E", and the nucleic acid sequence encoding it, which has potentially significant benefits for the treatment of tumours.

Therefore, according to a first aspect of the present invention there is provided a nucleic acid molecule encoding a VEGF-E protein or a functional equivalent, derivative or bioprecursor thereof, said protein comprising the amino acid sequence illustrated in Figure 2 or 4. Preferably, the nucleic acid molecule is a DNA and even more preferably a cDNA molecule.

Also provided by this aspect of the present invention is a nucleic acid molecule such as an antisense molecule capable of hybridising to the nucleic acid molecules according to the invention under high stringency conditions.

Stringency of hybridisation as used herein refers to conditions under which polynucleic acids are stable. The stability of hybrids is reflected in the melting temperature (T_m) of the hybrids. T_m can be approximated by the formula:

$$81.5^{\circ}\text{C} + 16.6(\log_{10}[\text{Na}^+] + 0.41 (\% \text{G\&C}) - 6001/l$$

wherein l is the length of the hybrids in nucleotides. T_m decreases approximately by 1-1.5°C with every 1% decrease in sequence homology.

The nucleic acid capable of hybridising to nucleic acid molecules according to the invention will generally be at least 70%, preferably at least 80 or 90% and more preferably at least 95% homologous to the nucleotide sequences according to the invention.

The present invention also comprises within its scope proteins or polypeptides encoded by the nucleic acid molecules according to the invention or a functional equivalent, derivative or bioprecursor thereof.

Therefore, according to a further aspect of the present invention, there is provided a VEGF-E protein, or a functional equivalent, derivative or bioprecursor thereof, having an amino acid sequence as illustrated in Figure 2 or 4. A further aspect of the invention comprises a VEGF-E protein, or a functional

equivalent, derivative or bioprecursor thereof,
encoded by a nucleic acid molecule according to the
invention. Preferably, the VEGF-E protein encoded by
said nucleic acid molecule comprises an amino acid
5 sequence as illustrated in Figure 2 or 4.

The DNA molecules according to the invention may,
advantageously, be included in a suitable expression
vector to express VEGF-E encoded therefrom in a
10 suitable host.

An expression vector according to the invention
includes a vector having a nucleic acid according to
the invention operably linked to regulatory sequences,
15 such as promoter regions, that are capable of
effecting expression of said DNA fragments. The term
"operably linked" refers to a juxta position wherein
the components described are in a relationship
permitting them to function in their intended manner.
20 Such vectors may be transformed into a suitable host
cell to provide for expression of a polypeptide
according to the invention. Thus, in a further
aspect, the invention provides a process for preparing
polypeptides according to the invention which
25 comprises cultivating a host cell, transformed or
transfected with an expression vector as described
above under conditions to provide for expression by
the vector of a coding sequence encoding the
polypeptides, and recovering the expressed
30 polypeptides.

The vectors may be, for example, plasmid, virus or
phage vectors provided with an origin of replication,
optionally a promoter for the expression of said
35 nucleotide and optionally a regulator of the promoter.

The vectors may contain one or more selectable markers, such as, for example, ampicillin resistance.

5 Regulatory elements required for expression include promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. For example, a bacterial expression vector may include a promoter such as the lac promoter and for transcription initiation the Shine-Dalgarno
10 sequence and the start codon AUG. Similarly, a eukaryotic expression vector may include a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of
15 the ribosome. Such vectors may be obtained commercially or assembled from the sequences described by methods well known in the art.

20 Nucleic acid molecules according to the invention may be inserted into the vectors described in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense nucleic acids may be produced by synthetic means.

25 In accordance with the present invention, a defined nucleic acid includes not only the identical nucleic acid but also any minor base variations including in particular, substitutions in bases which result in a synonymous codon (a different codon specifying the
30 same amino acid residue) due to the degenerate code in conservative amino acid substitutions. The term "nucleic acid sequence" also includes the complementary sequence to any single stranded sequence given regarding base variations.

The present invention also advantageously provides nucleic acid sequences of at least approximately 10 contiguous nucleotides of a nucleic acid according to the invention and preferably from 10 to 50

5 nucleotides. These sequences may, advantageously be used as probes or primers to initiate replication, or the like. Such nucleic acid sequences may be produced according to techniques well known in the art, such as by recombinant or synthetic means. They may also be

10 used in diagnostic kits or the like for detecting the presence of a nucleic acid according to the invention. These tests generally comprise contacting the probe with the sample under hybridising conditions and detecting for the presence of any duplex or triplex

15 formation between the probe and any nucleic acid in the sample.

The nucleic acid sequences according to this aspect of the present invention comprises the sequences of

20 nucleotides designated herein as VEGFE 1-10, illustrated in Figure 5.

According to the present invention these probes may be anchored to a solid support. Preferably, they are

25 present on an array so that multiple probes can simultaneously hybridize to a single biological sample. The probes can be spotted onto the array or synthesised *in situ* on the array. (See Lockhart et al., Nature Biotechnology, vol. 14, December 1996

30 "Expression monitoring by hybridisation to high density oligonucleotide arrays". A single array can contain more than 100, 500 or even 1,000 different probes in discrete locations.

35 The nucleic acid sequences, according to the invention

may be produced using such recombinant or synthetic means, such as for example using PCR cloning mechanisms which generally involve making a pair of primers, which may be from approximately 10 to 50 nucleotides to a region of the gene which is desired to be cloned, bringing the primers into contact with mRNA, cDNA, or genomic DNA from a human cell, performing a polymerase chain reaction under conditions which bring about amplification of the desired region, isolating the amplified region or fragment and recovering the amplified DNA. Generally, such techniques as defined herein are well known in the art, such as described in Sambrook et al (Molecular Cloning: a Laboratory Manual, 1989).

The nucleic acids or oligonucleotides according to the invention may carry a revealing label. Suitable labels include radioisotopes such as ^{32}P or ^{35}S , enzyme labels or other protein labels such as biotin or fluorescent markers. Such labels may be added to the nucleic acids or oligonucleotides of the invention and may be detected using known techniques *per se*.

The protein according to the invention includes all possible amino acid variants encoded by the nucleic acid molecule according to the invention including a polypeptide encoded by said molecule and having conservative amino acid changes. Proteins or polypeptides according to the invention further include variants of such sequences, including naturally occurring allelic variants which are substantially homologous to said proteins or polypeptides. In this context, substantial homology is regarded as a sequence which has at least 70%, preferably 80 or 90% amino acid homology with the

proteins or polypeptides encoded by the nucleic acid molecules according to the invention.

5 The nucleic acid or protein according to the invention may be used as a medicament or in the preparation of a medicament for treating cancer or other diseases or conditions associated with expression of VEGF-E protein.

10 Advantageously, the nucleic acid molecule or the protein according to the invention may be provided in a pharmaceutical composition together with a pharmacologically acceptable carrier, diluent or excipient therefor.

15 The present invention is further directed to inhibiting VEGF2 *in vivo* by the use of antisense technology. Antisense technology can be used to control gene expression through triple-helix formation
20 or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the mature protein sequence, which encodes for the protein of the present invention, is used to design an antisense RNA
25 oligonucleotide of from 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple-helix - see Lee *et al.* Nucl. Acids Res., 6:3073 (1979); Cooney *et al.*, Science, 241:456 (1988);
30 and Dervan *et al.*, Science, 251: 1360 (1991), thereby preventing transcription and the production of VEGF2. The antisense RNA oligonucleotide hybridises to the mRNA *in vivo* and blocks translation of an mRNA molecule into the VEGF2 (antisense - Okano, J.
35 Neurochem., 56:560 (1991); Oligodeoxynucleotides as

Antisense Inhibitors of Gene Expression, CRC Press,
Boca Raton, FL (1988)).

5 Alternatively, the oligonucleotide described above can
be delivered to cells by procedures in the art such
that the anti-sense RNA or DNA may be expressed *in*
vivo to inhibit production of VEGF-E in the manner
described above.

10 Antisense constructs to VEGF-E, therefore, may inhibit
the angiogenic activity of the VEGF-E and prevent the
further growth or even regress solid tumours, since
angiogenesis and neovascularization are essential
15 steps in solid tumour growth. These antisense
constructs may also be used to treat rheumatoid
arthritis, psoriasis and diabetic retinopathy which
are all characterized by abnormal angiogenesis.

20 A further aspect of the invention provides a host cell
or organism, transformed or transfected with an
expression vector according to the invention. The
host cell or organism may advantageously be used in a
method of producing VEGF-E, which comprises recovering
any expressed VEGF-E from the host or organism
25 transformed or transfected with the expression vector.

According to a further aspect of the invention there
is also provided a transgenic cell, tissue or organism
comprising a transgene capable of expressing VEGF-E
30 protein according to the invention. The term
"transgene capable of expression" as used herein means
a suitable nucleic acid sequence which leads to
expression of VEGF-E or proteins having the same
function and/or activity. The transgene, may include,
35 for example, genomic nucleic acid isolated from human

cells or synthetic nucleic acid, including DNA integrated into the genome or in an extrachromosomal state. Preferably, the transgene comprises the nucleic acid sequence encoding the proteins according to the invention as described herein, or a functional fragment of said nucleic acid. A functional fragment of said nucleic acid should be taken to mean a fragment of the gene comprising said nucleic acid coding for the proteins according to the invention or a functional equivalent, derivative or a non-functional derivative such as a dominant negative mutant, or bioprecursor of said proteins. For example, it would be readily apparent to persons skilled in the art that nucleotide substitutions or deletions may be used using routine techniques, which do not affect the protein sequence encoded by said nucleic acid, or which encode a functional protein according to the invention.

VEGF-E protein expressed by said transgenic cell, tissue or organism or a functional equivalent or bioprecursor of said protein also form part of the present invention.

Antibodies to the protein or polypeptide of the present invention may, advantageously, be prepared by techniques which are known in the art. For example, polyclonal antibodies may be prepared by inoculating a host animal, such as a mouse, with the polypeptide according to the invention or an epitope thereof and recovering immune serum. Monoclonal antibodies may be prepared according to known techniques such as described by Kohler R. and Milstein C., Nature (1975) 256, 495-497.

Antibodies according to the invention may also be used in a method of detecting for the presence of a polypeptide according to the invention, which method comprises reacting the antibody with a sample and
5 identifying any protein bound to said antibody. A kit may also be provided for performing said method which comprises an antibody according to the invention and means for reacting the antibody with said sample.

10 Proteins which interact with the polypeptide of the invention may be identified by investigating protein-protein interactions using the two-hybrid vector system first proposed by Chien et al (1991).

15 This technique is based on functional reconstitution in vivo of a transcription factor which activates a reporter gene. More particularly the technique comprises providing an appropriate host cell with a DNA construct comprising a reporter gene under the
20 control of a promoter regulated by a transcription factor having a DNA binding domain and an activating domain, expressing in the host cell a first hybrid DNA sequence encoding a first fusion of a fragment or all of a nucleic acid sequence according to the invention
25 and either said DNA binding domain or said activating domain of the transcription factor, expressing in the host at least one second hybrid DNA sequence, such as a library or the like, encoding putative binding proteins to be investigated together with the DNA
30 binding or activating domain of the transcription factor which is not incorporated in the first fusion; detecting any binding of the proteins to be investigated with a protein according to the invention by detecting for the presence of any reporter gene
35 product in the host cell; optionally isolating second

hybrid DNA sequences encoding the binding protein.

An example of such a technique utilises the GAL4 protein in yeast. GAL4 is a transcriptional activator of galactose metabolism in yeast and has a separate domain for binding to activators upstream of the galactose metabolising genes as well as a protein binding domain. Nucleotide vectors may be constructed, one of which comprises the nucleotide residues encoding the DNA binding domain of GAL4. These binding domain residues may be fused to a known protein encoding sequence, such as for example the nucleic acids according to the invention. The other vector comprises the residues encoding the protein binding domain of GAL4. These residues are fused to residues encoding a test protein. Any interaction between polypeptides encoded by the nucleic acid according to the invention and the protein to be tested leads to transcriptional activation of a reporter molecule in a GAL-4 transcription deficient yeast cell into which the vectors have been transformed. Preferably, a reporter molecule such as β -galactosidase is activated upon restoration of transcription of the yeast galactose metabolism genes.

Advantageously, the antibody according to the invention may also be used as a medicament or in the preparation of a medicament for treating tumours or other diseases associated with expression of VEGF-E. The invention also further provides a pharmaceutical composition comprising said antibody together with a pharmaceutically acceptable carrier diluent or excipient therefor.

A further aspect of the present invention also

provides a method of identifying VEGF-E in a sample, which method comprises contacting said sample with an antibody according to the invention and monitoring for any hybridisation of any proteins to said antibody. A
5 kit for identifying the presence of VEGF-E in a sample is also provided comprising an antibody according to the invention and means for contacting said antibody with said sample.

10 The invention may be more clearly understood with reference to the accompanying example, which is purely exemplary, with reference to the accompanying drawings, wherein:

15 Figure 1: is a nucleotide sequence coding for a partial VEGF-E protein according to the invention.

20 Figure 2: is an illustration of amino acid sequence of the nucleic acid sequence of Figure 1.

Figure 3: is an illustration of a nucleotide sequence encoding VEGF-E protein according to the invention.

25 Figure 4: is an illustration of the amino acid sequence of the nucleic acid sequence of Figure 3.

30 Figure 5: depicts the nucleic acid sequences of the first 18 human EST clones obtained from the BLAST search of the LifSeqTM database.

35 Figure 6: depicts the nucleotide sequences of 50 human EST clones obtained from the proprietary

LifeSeq™ database.

Figure 7: is an illustration of the nucleotide
sequences utilised as primers to identify
5 the sequence of the gene coding for VEGF-E.

EXAMPLE 1

A BLAST (Basic Local Alignment Search Tool; Altschul
10 et al., 1990 J. Mol. Biol. 215, 403-410) search was
performed in the propriety LifeSeq™ human EST database
(Incyte Pharmaceuticals, Inc., Palo Alto, CA, USA).
BLAST produces alignments of both nucleotide and amino
acid sequences to determine sequence similarity.
15 Because of the local nature of the alignments, BLAST
is especially useful in determining exact matches or
in identifying homologues. While it is useful for
matches which do not contain gaps, it is inappropriate
for performing motif-style searching. The fundamental
20 unit of BLAST algorithm output is the High-scoring
Segment Pair (HSP).

Eighteen human EST clones (Figure 5) with high
similarity to the previously identified VEGF proteins
25 were identified and a further fifty EST clones (Figure
6) were identified using these sequences as query
sequences, allowing us to deduce the putative sequence
for the new VEGF-E protein. The sequences obtained
were compared to known sequences to determine regions
30 of homology and to identify the sequence as a novel
VEGF-E protein. Using the DNA sequence information in
the databases we were able to prepare suitable primers
having the sequences of VEGFE 1-10 illustrated in
Figure 7 for use in subsequent RACE experiments to
35 obtain the complete DNA sequence for the VEGF-E gene.

CLAIMS

1. A nucleic acid molecule encoding a VEGF-E protein
or a functional equivalent derivative or bioprecursor
5 thereof, said protein comprising the amino acid
sequence illustrated in Figures 2 or 4.
2. A nucleic acid molecule according to claim 1
wherein said nucleic acid is a DNA molecule.
10
3. A nucleic acid molecule according to claim 1 or 2
wherein said nucleic acid is a cDNA molecule.
4. A nucleic acid molecule according to any of
15 claims 1 to 3 comprising the nucleotide sequence
illustrated in Figure 1 or 3.
5. A nucleic acid molecule capable of hybridising to
a molecule according to any of claims 1 to 4 under
20 high stringency conditions.
6. A VEGF-E protein, or a functional equivalent,
derivative or bioprecursor thereof, having the amino
acid sequence illustrated in Figure 2 or 4.
25
7. A VEGF-E protein, or a functional equivalent,
derivative or bioprecursor thereof, encoded by a
nucleic acid molecule according to any of claims 1 to
4.
30
8. A protein according to claim 7, which comprises
the amino acid sequence illustrated in Figure 2 or 4.
9. An expression vector comprising a nucleic acid
35 molecule according to any of claims 1 to 4.

10. An expression vector according to claim 9 further comprising a nucleotide sequence encoding a reporter molecule.

5 11. A nucleic acid molecule according to any of claims 1 to 5 for use as a medicament.

12. Use of a nucleic acid molecule according to any of claims 1 to 5 in the preparation of a medicament
10 for inhibiting angiogenic activity and formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair or for treating cancer or rheumatoid arthritis or psoriasis or diabetic retinopathy.

15 13. A pharmaceutical composition comprising a nucleic acid molecule or a protein according to any of claims 1 to 5 or 6 to 8 respectively, together with a pharmaceutically acceptable carrier, diluent or
20 excipient therefor.

14. A host cell or organism transformed or transfected with an expression vector according to claim 9 or 10.

25 15. A transgenic cell, tissue or organism comprising a transgene capable of expressing a VEGF-E protein according to any of claims 6 to 8.

30 16. A process for producing a VEGF-E protein according to any of claims 6 to 8, said process comprising transforming a host cell or organism with an expression vector according to claim 9 and 10, and recovering the expressed protein from said host cell
35 or organism.

17. An antibody capable of binding to a protein according to any of claims 6 to 8, which is preferably a monoclonal antibody.

5 18. An antibody according to claim 17 for use as a medicament.

10 19. Use of an antibody according to claim 17 in the preparation of a medicament for inhibiting angiogenic activity and formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair or for treating cancer or rheumatoid arthritis or psoriasis or diabetic retinopathy.

15

20 20. A pharmaceutical composition comprising an antibody according to claim 17 together with a pharmaceutically acceptable carrier diluent or excipient therefor.

25 21. A method of identifying VEGF-E in a sample which method comprises contacting said sample with an antibody according to claim 17 and monitoring for binding of any protein to said antibody.

30 22. A kit for identifying the presence of VEGF-E in a sample which comprises an antibody according to claim 17 and means for contacting said antibody with said sample.

35 23. A method of identifying compounds which inhibit angiogenesis which method comprises providing a host cell or organism according to claim 14 or a transgenic

cell, tissue or organism according to claim 15,
contacting a test compound with said cell, tissue or
organism and monitoring for the presence or absence
either of said reporter molecule or VEGF-E.

5

24. A compound identifiable according to the method
of claim 23.

25. A compound according to claim 24 for use as a
10 medicament.

26. Use of a compound according to claim 24 in the
preparation of a medicament for inhibiting angiogenic
activity and formation and proliferation of new blood
15 vessels, growth and development of tissues, tissue
regeneration and organ and tissue repair or for
treating cancer, rheumatoid arthritis, psoriasis or
diabetic retinopathy.

20 27. A nucleic acid sequence comprising the nucleotide
sequence of any of the sequences identified in Figure
6 or 7.

28. An expression vector comprising a nucleic acid
25 sequence according to claim 27.

29. A host cell transformed or transfected with an
expression vector according to claim 28.

30 30. A method for producing a polypeptide, said method
comprising the steps of:

- a) culturing the host cell of claim 29 under
conditions suitable for expression of the
peptide; and
- 35 b) recovering the polypeptide from the host
cell culture.

+3 M N I F L L N L L T E E V R L Y
}-----
1 AGGAAATCAA ATTAGATAA GATTGTATC TGATGAATAT TTCTCTCTG AACCTTCTAA CAGAGGAGGT AAGATTATAC
TCCTTTAGTT TAATCCTATT CTAACATAG ACTACTTATA AAAGGAAGAC TTGGAAGATT GTCTCCTCCA TTCTAATATG
.....
+3 S C T P R N P S V S I R E E L K R T D T I F W P G C L
}-----
81 AGCTGCACAC CTCGTAACCT CTCAGTGTCC ATAAGCGAAG AACTAAAGAG AACCGATACC ATTTCTGCGC CAGGTTGTCT
TCGACGTGTG GAGCATTGAA GAGTCACAGG TATTCCTTTC TTGATTCTC TTGGCTATGG TAAAAGACCG GTCCAACAGA
-2 <-----
.....
+3 L V K R C G G N C A C C L H N C N E C Q C V P S K V
}-----
161 CCTGGTTAAA CGCTGTGGTG GGAACGTGTC CTGTTGTCTC CACAATTGCA ATGAATGTCA ATGTGTCCCA AGCAAAGTTA
GGACCAATTT GCGACACCAC CCTTGACACG GACAACAGAG GTGTTAAGT TACTTACAGT TACACAGGCT TCGTTTCAAT
-2 -----
.....
+3 T K K Y H E V L Q L R P K T G V R G L H K S L T D V A
}-----
+1 V S G D C T N H S P T W P
}-----
241 CTAAAAATA CCACGAGGTC CTTCAGTTGA GACCAAGAC CGGTGTACG GGATTGCACA AATCACTCAC CGACGTGGCC
GATTTTTTAT GGTGCTCCAG GAAGTCAACT CTGTTTCTG GCCACAGTCC CTAACGTGT TTAGTGAGTG GCTGCACCGG
-2 -----
.....
+3 L E H H E E C D C V C R G S T G G
}-----
+2 V Q R E H R R I A A S P P A A L A
}-----
+1 W S T M R S V T V C A E G A Q E D S R I T T S S S C
}-----
321 CTGGAGCACC ATGAGGAGTG TGACTGTGTG TGCAGAGGGA GCACAGGAGG ATAGCCGCAT CACCACCAGC AGCTCTTGCC
GACCTCGTGG TACTCCTCAC ACTGACACAC ACCTCTCCCT CGTGTCTCTC TATCGGCGTA GTGGTGGTCG TCGAGAACCG
.....
+2 Q S C A V Q W L I L L E N V C V I S I L N L S C L L Q
}-----
+1 P E L C S A V A D S I R E R M R Y L H P
}-----
401 CAGAGCTGTG CAGTGCAGTG GCTGATTCTA TTAGAGAACG TATGCGTTAT CTCCATCCTT AATCTCAGTT GTTGTCTTCA
GTCTCGACAC GTCACGTCAC CGACTAAGAT AATCTCTTGC ATACGCAATA CAGGTAGGAA TTAGAGTCAA CAAACGAAGT
.....
+2 G P F I F R I Y S A F
}-----
481 AGGACCTTTC ATCTTCAGGA TTTACAGTGC ATTCTGAAAG AGGAGACATC AAACAGAATT AGGAGTTGTG CAACAGCTCT
TCCTGGAAAG TAGAAGTCCT AAATGTCACG TAAGACTTTC TCCTCTGTAG TTGTCTTAA TCCTCAACAC GTTGTGAGA
.....
561 TTTGAGAGGA GGCCTAAAGG ACAGGAGAAA AGGTCTTCAA TCGTGGAAAG AAAATTAAAT GTGTATTAA ATAGATCACC
AAACTCTCCT CCGGATTTC TGTCCCTTT TCCAGAAGT AGCACCTTTC TTTTAATTA CAACATAATT TATCTAGTGG
.....
641 AGCTAGTTTC AGAGTTACCA TGTACGTATT CCACTAGCTG GGTCTGTAT TTCAGTTCTT TCGATACGGC TTAGGTAAT
TCGATCAAAG TCTCAATGGT ACATGCATAA GGTGATCGAC CCAAGACATA AAGTCAAGAA AGCTATGCCG AATCCATTA
.....
721 GTCAGTACAG GAAAAAACT GTGCAAGTGA GCACCTGATT CGGTGCTT GGCTTAACTC TAAAGCTCCA TGTCTGGGC
CAGTCATGTC CTTTTTTTGA CACGTTCACT CGTGGACTAA GGCAACGGAA CCGAATTGAG ATTTGAGGT ACAGGACCCG
.....
801 CTAAATCGT ATAAATCTG GA
GATTTTAGCA TATTTTAGAC CT

1 MNIFLLNLLT EEVRLYSCTP RNFSVSIREE LKRTDTIFWP GCLLVKRCGG
.....
51 NCACCLHNCTN ECQCVPKVT KKYHEVLQLR PKTGVRGLHK SLTDVALEHH
.....
101 EECDCVCRGS TGG
.....

Fig 2

+3 M N I F L L N L L T E E V R L Y
1 AGGAAATCAA ATTAGGATAA GATTGTATC TGATGAATAT TTTCCTTCTG AACCTTCTAA CAGAGGAGGT AAGATTATAC
TCCTTTAGTT TAATCCTATT CTAACATAG ACTACTTATA AAAGGAAGAC TTGGAAGATT GTCTCCTCCA TTCTAATATG
+3 S C T P R N F S V S I R E E L K R T D T I F W P G C L
81 AGCTGCACAC CTCGTAACCT CTCAGTGTCC ATAAGGGAAG AACTAAAGAG AACCGATACC ATTTTCTGGC CAGGTTGTCT
TCGACGTGTG GAGCATTGAA GAGTCACAGG TATTCCCTTC TTGATTCTC TTGGCTATGG TAAAGACCG GTCCAACAGA
-2
+3 L V K R C G G N C A C C L H N C N E C Q C V P S E V
161 CCTGGTTAAA CGCTGTGGTG GGAACGTGTC CTGTTGTCTC CACAATTGCA ATGAATGTCA ATGTGTCCCA ASCAAAGTA
GGACCAATTT GCGACACCAC CCTTGACACG GACAACAGAG GTGTTAACGT TACTTACAGT TACACAGGGT TCGTTTCAAT
-2
+3 T K K Y H E V L Q L R P K T G V R G L H K S L T D V A
+1 V S G D C T N H S P T W P
241 CTAATAAATA CCACGAGGTC CTTCACTTGA GACCAAGAC CGGTGTGAGG GGATTGCACA AATCACTCAC CGACGTGGCC
GATTTTAT GGTGCTCCAG GAAGTCAACT CTGGTTCTG GCCACAGTCC CTAACGTGT TTAGTGAGTG CCTGCACCGG
-2
+3 L E H H E E C D C V C R G S T G G
+2 V Q R E H R R I A A S P P A A L A
+1 W S T M R S V T V C A E G A Q E D S R I T T S S S C
321 CTGGAGCACC ATGAGGAGTG TGACTGTGTG TGCAGAGGGA GCACAGGAGG ATAGCCGCAT CACCACCAGC AGCTCTTGGC
GACCTCGTGG TACTCCTCAC ACTGACACAC ACGTCTCCTT CGTGTCTTCC TATCGGCGTA GTGGTGGTGG TCGAGAACGG
+2 Q S C A V Q W L I L L E N V C V I S I L N L S C L L Q
+1 P E L C S A V A D S I R E R M R Y L H P
401 CAGAGCTGTG CAGTGCAGTG GCTGATTCTA TTAGAGAACG TAIGCGTTAT CTCCATCCTT AATCTCAGTT GTTGCTTCA
GTCTCGACAC GTCACGTCAC CGACTAAGAT AATCTCTTGC ATACGCAATA GAGGTAGGAA TTAGAGTCAA CAAACGAAGT
+2 G P F I F R I Y S A F
481 AGGACCTTTC ATCTTCAGGA TTTACAGTGC ATTCTGAAAG AGGAGACATC AAACAGAATT AGGAGTTGTG CACACGCTCT
TCCTGGAAAG TAGAAGTCCT AAATGTCACG TAAGACTTTC TCCTCTGTAG TTGTCTTAA TCCTCAACAC GTTGTGAGA
561 TTTGAGAGGA GGCCTAAAGG ACACGAGAAA AGGCTCTTCA TCGTGCAAAG AAAATTAAAT GTTGATTAA ATAGATCACC
AAACTCTCCT CCGGATTTC TGTCTCTTT TCCAGAAGT AGCACCTTTC TTTTAATTTA CAACATAATT TATCTAGTGG
641 AGCTAGTTTC AGAGTTACCA TGTACGTATT CCACTAGCTG GGTCTGTAT TTCAGTCTT TCGATACGGC TTAGGGTAAT
TCGATCAAAG TCTCAATGGT ACATGCATAA GGTGATCGAC CCAAGACATA AAGTCAAGAA AGCTATGCCG AATCCCATTA
721 GTCAGTACAG GAAAAAACT GTGCAAGTGA GCACCTGATT CGGTGCTT GGCTTAACTC TAAAGCTCCA TGTCTCGGC
CAGTCATGTC CTTTTTTTGA CAGTTTCACT CGTGGACTAA GGCAACGGAA CCGAATTGAG ATTTGAGGT ACAGGACCCG
801 CTAATAATCGT ATAAATCTG GATTTTTTNN TTTTTTTTG CGCATATTCA CATATGTAAA CCAGAACCT CTATGTACTA
GATTTTAGCA TATTTTAGAC CTAATAAAN AAAAAAAC CGGTATAAGT GTATACATTT GGTCTGTAA GATACATGAT
881 CAAACCTGGT TTTTAAAAAG GAACTATGTT GCTATGAATT AAAGTTGTTT CGTGTGATA GGACAGACTG GATTTTTCAT
GTTTGGACCA AAAATTTTTC CTGATACAA CGATACTTAA TTGAACACA GCACGACTAT CTTGTCTGAC CTAATAAGTA
-3

Fig 3

961 ATTTCTTATT AAAATTTCTG CCATTAGAA GAAGAGAACT ACATTCATCG TTGGAAGAG ATAAACCTGA AAGAAGACT
TAAAGAATAA TTTAAAGAC GGTAAATCTT CTTCTCTTGA TGTAAGTACC AAACCTTCTC TATTGGACT TTTCTCTCA
-3 -----
1041 GGCCTTATCT TCACTTTATC GATAAGTCAG TTTATTTGTT TCATTGTGTA CATTTTATA TTCTCCTTT GACATTATAA
CCGGAATAGA AGTGAAATAG CTATTCAGTC AAATAAACAA AGTAACACAT GTAAATATAT AAGAGGAAAA CTGTAATATT
-3 -----
1121 CTGTTGGCTT TTCTAATCTT GTTAAATATA TCTATTTTGA CCAAAGGTAT TTAATATTCT TTTTATGAC AACTTAGATC
GACAACCGAA AAGATTAGAA CAATTATAT AGATAAAAT GGTTCCTATA AATTATAAGA AAAAATACTG TTGAATCTAG
1201 AACTATTTT AGCTTGGTAA ATTTTCTAA ACACAATTGT TATAGCCAGA GGAACAAAGA TGATATAAAA TATTGTTGCT
TTGATAAAAA TCGAACCATT TAAAAAGATT TGTGTTAACA ATATCGGTCT CCTTGTTTCT ACTATATTT ATAACAACGA
1281 CTGACAAAAA TACATGTATT TCATCTCTGT ATGGTGTAG AGTTAGATTA ATCTGCATTT TAAAAAACTG AATTGGAATA
GACTGTTTTT ATGTACATAA AGTAAGAGCA TACCACGATC TCAATCTAAT TAGACGTAAA ATTTTTTGAC TTAACCTTAT
1361 GAATTGGTAA GTTGCAAAGA CTTTTTGAAA ATAATTAAAT TATCATATCT TCCATTCCTG TTATTGGAGA TGAAAAATAA
CTTAACCATT CAACGTTTCT GAAAAACTTT TATTAATTTA ATAGTATAGA AGGTAAGGAC AATAACCTCT ACTTTTATTT
1441 AAGCAACTTA TGAAAGTAGA CATTGAGATC CAGCCATTAC TAACCTATTC CTTTTTGGG GAAATCTGAG CCTAGCTCAG
TTCGTTGAAT ACTTTCATCT GTAAGTCTAG GTCGGTAATG ATTGGATAAG GAAAAAACC CTTTAGACTC GGATCGAGTC
1521 AAAACATAA AGCACCTTGA AAAAGACTTG GCAGCTTCCT GATAAAGCGT GCTGTGCTGT GCAGTAGGAA CACATCCTAT
TTTTGTATT TCGTGGAAGT TTTCTGAAC CGTCGAAGCA CTATTTGCA CGACACGACA CGTCATCCTT GTGTAGGATA
1601 TTATTGTGAT GTTGTGGTTT TATTATCTTA AACTCTGTTT CATACTTGT TATAAATACA TGGATATTTT TATGTACAGA
AATAACACTA CAACACCAAA ATAATAGAAT TTGAGACAAG GTATGTGAAC ATATTTATGT ACCTATAAAA ATACATGTCT
1681 AGTATGTCTC TTAACCAGTT CACTATTGT ACCTGG
TCATACAGAG AATTGGTCAA GTGAATAACA TGGACC

Fig 3 (cont'd)

1 MNIFLLNLLT EEVRLYSCTP RNFVSISIREE LKRTDTISWF GCLLVKRCGG
.....
51 NCACCLHNCN ECQCVPKVT KKYHEVLQLR PKTGVRGLHK SLTUVALEHH
.....
101 EECDCVCRGS TGG
.....

Figure 4

1 >3993180H1 LUNGNON01 INCYTE
 2 CACAP...CACTCACCGACGTGGCCCTGGAGCACCATGAGGNGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCC
 3 GCAT...CACCAGCAGTCTTTGCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCAT
 4 CCTTAATCTCAGTTGTTTGTCTCAAGGACCTTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAO
 5 AATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCTAAAGGACAGGAGAAAGGTCTT
 6 >3510192H1 CONCN01 INCYTE
 7 TGCAGTCCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGTCTCAAGGACCTT
 8 TCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAG
 9 GAGCCCTAAAGGACAGGAGAAAGGTCTTCAATCGTGGAAAGAAATTAATGTTGTATTAAATAGATCACCAGCTAGTT
 10 TCAGAGTTACCATGTACGTATTCCACTAGCTGGCTTCTGTATT
 11 >2559870H1 ADRETUT01 INCYTE
 12 CACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAAGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCA
 13 TGAGGAGTGTGACTGTGTGTCAGAGGGAGCACAGGGGGATAGCCGCATCACCACCAGCAGCTCTTGCCAGAGCTGTGC
 14 AGTGCAGTGGCTGACTTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGTCTCAAGGACCTTTCA
 15 TCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGA
 16 >3979767H1 LUNGTUT08 INCYTE
 17 GGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGC
 18 GTTATCTCCATCCTTAATCTCAGTTGTTTGTCTCAAGGACCTTTTCATCTCAGGATTTACAGTGCATTCTGAAAGAGGAG
 19 ACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCTAAAGGACAGGAGAAAGGTCTTCAATCGTG
 20 GAAAGAAATTAATGTTGTATTAAATAGACACCAGCT
 21 >3980011H1 LUNGTUT08 INCYTE
 22 GGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGC
 23 GTTATCTCCATCCTTAATCTCAGTTGTTTGTCTCAAGGACCTTTTCATCTCAGGATTTACATGCATTCTGAAAGAGGAGA
 24 CATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCTAAAGGACAGGAGAAAGGTCTTCAATCGTG
 25 AAAGAAATTAATGTTGTATTAAATAGATCACCA
 26 >4825396H1 BLADDIT01 INCYTE
 27 GGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAAGTGTGCTGTGTCTCCACAATT
 28 GCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTC
 29 AGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCATGAGGAGTGTCACTGTGTGTGCAGAGGGAGCACAGG
 30 AGGATAGCCGCATCACCACCA
 31 >3073703H1 BONEUNT01 INCYTE
 32 AGAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACCTCTCAGT
 33 GTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAAGT
 34 GTGCTGTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAG
 35 TTGAGACCAAAGACCGGTGTGTCAGGGGATTGCACAAATCA
 36 >1302516H1 PLACNOT02 INCYTE
 37 AGGAAATCAAATTAGGATAAGATTGTATCTGATGAATATTTTCTCTGAACCTTCTAACAGAGGAGGTAAGATTATAC
 38 AGCTGCACACCTCGTAACCTTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGCCCAGGTTGTCT
 39 CCTGGTTAAACGCTGTGGTGGGAAGTGTGCTGTGTCTCCACAATTGCAATGAATGTCAATGTGTGTCCCAAGCAAGTT
 40 ACTAAAAATACCACGAGGTC
 41 >3684109H1 HEANOT01 INCYTE
 42 ATTTTCATCTTCAGGATTTACAGTGCATTCTGAAANAGGAGAAATCAAACANAATTAGGAGTTGTGCAACAGCTCTTTTGA
 43 GAGGAGGCCTAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAANAAAATTAATGTTGTATTAAATAGATCACCAGCTA
 44 GTTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTTCGATACGGCTTAGGGTAATGTGAG
 45 TACAGGAAAAAAGTGTGCAAGTGAGCACCTGATTCCTGCTTGCCTTGCCTT
 46 >4713188H1 BRAIHCT01 INCYTE
 47 CAAAGTTACTAAAAAATACCACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTGTCAGGGGATTGCACAAATCACTCACCG
 48 ACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAG
 49 CTCTTGCCAGAGCTGTGTCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGT
 50 TTGCT
 51 >458823H1 KERANOT01 INCYTE
 52 ANGAGTTGCCCAGAGCTGTGTCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTT
 53 GTTTGNTTCAAGGACCTTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTG
 54 CAACAGCTCTTTTGAGAGGAGGCTAAAGGACAGGAGAAAGGTCTTCAATCGTGGAAAGAAAATTAATGTTGTATTAA
 55 ATAGATC
 56 >1303909H1 PLACNOT02 INCYTE
 57 AGGAAATCAAATTAGGATAAGATTGTATCTGATGAATATTTTCTCTGAACCTTCTAACAGAGGAGGTAAGATTATAC
 58 AGCTGCACACCTCGTAACCTTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGCCCAGGTTGTCT
 59 CCTGGTTAAACGCTGTGGTGGGAAGTGTGCTGTCTCCCAATGCAATGAATGTCAATGTGTGTCCCAAG
 60 >2739211H1 OVARNOT09 INCYTE
 61 GTGCTTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCTAAAGGACAGGA
 62 GAAAAGGTCTTCAATCGTGGAAAGAAAATTAATGTTGTATTAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACG
 63 TATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTTCGATACGGCTTAGGGTAATGTGTCAGTACAGGAAAAAACTGTGCAA
 64 GTGAGCACCTGAT
 65 >3325591H1 PTHYNOT03 INCYTE
 66 TGCAACAGCTCTTTTGAGAGGAGGCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAATGTTGTATT
 67 AAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTTCGATACG
 68 GCTTAGGGTAATGTGTCAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCTTGCCTTGCCTTAACCTTAAAGCNCC
 69 ATGTCNNNGGCNAAAANCAGAAAAAT
 70 >3733565H1 SMCCNOS01 INCYTE
 71 CCTTAATCTCAGTTGTTTGTCTCAAGGACCTTTTCATCTTCAGGATTTACAGTGCATTCTGNAAGANGAGACATCAAACAG
 72 AATTAGGNGTTGTGCAAAAGCTCTTTTGAGAGGAGGCTAAAGGACAGGAGAAAAGGTCTNCAATCGTGGAAAGNAAAAT
 73 AAATCTGTATNAATNGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGNCNGTATTACAGTCT
 74 TTCGGAACGGCTTAGGGTAATGTGTCAGTACAGGANAAAACTGTGTCAGTGAG
 75 >3554223H1 SYNONOT01 INCYTE

file: Short-est.mfcs

76 ATTAAATAGATCACCAGCTAGTTTACAGATTACCATGTACGTATTCCTAGCTGGGCTGTATTTCAGTTCTTTGGAT
 77 ACGGCTTAGGGTAATGTCAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGGCTTAACTCTAAAG
 78 CTCCTTCCTGGGCCTAAAATCGTATAAAATCTGGATTTTTTTTNTTTTTTTTTTGGCATATTCACATATGTAAACCAGN
 79 ACATTCTATGTACNACAAACCTGGTTTTTAAAAAGGAAC
 80 >4507477H1 OVARTDT01 INCYTE
 81 GGTAGTTTCAGAGTTACCATGTACGTATTCCTAGCTGGGCTGTATTTCAGTTCTTTGGATACGGCTTAGGCTAAT
 82 GTCAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGGCTTAACTCTAAAGCTCCATGTCTGSGCC
 83 TAAATCGTATAAAATCTGGA
 84 >4163378H1 BRSTNOT32 INCYTE
 85 AATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCTAGCTGGGNTCTGTATTTCAGTTCTTTGGATACG
 86 GCTTAGGGTAATGTCAGTACAGGAAAAAGCTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGGCTTAACTCTAAGCTCC
 87 ATGTCCTGGGCTAAAATCTATA

Fig 5(cont'd)

1 >2054675H1 BEPINOT01 INCYTE
2 AAAGGACCTATGTTGCTATGAATTAACCTTGTGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAATAATT
3 TCTGCTTTAGAGAAGAGACAACATTCATTTGTTGGAAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTT
4 TATCGATAAGCTCAGTTTATTTGTTTCATTTGTGTACATTTTATATTTCTCTTTTGACATTATAACTGTTGGCTTTTCTAA
5 TCTTGTAAATATATCTATTTTACCAGGATTTTAAATATTTCTTTTAA
6 >3993180H1 LUNGNON03 INCYTE
7 CACAAATCACTCACCGACGTGGCCCTGGAGCACCATTGAGGNGTGTGACTGTGTGTCAGAGGGAGCACAGGAGGATAGCC
8 GCATCACCACCAGCAGCTCTTGGCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCAT
9 CCTTAATCTCAGTTGTTTGTCTCAAGGACCTTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAG
10 AATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCTAAAGGACAGGAGAAAGGTCCT
11 >3510192H1 CONCNOT01 INCYTE
12 TGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGTCTCAAGGACCTT
13 TCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAG
14 GAGGCCATAAGGACAGGAGAGAAAGGTCTTCAATCGTGGAAAGAAATTAATGTTGTATTAAATAGATCACCAGCTAGTT
15 TCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTT
16 >4164633H1 BRSTNOT12 INCYTE
17 CTGTGTTAAATATATCTATTTTACCAAAGGTATTTAATATCTTTANTTTATGACAACCTAGATCAACTATTTTGTAGCTTG
18 GTAAATTTTCTTAAACACAAATGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATG
19 TATTTTCATCTCGTATGGTGTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCA
20 AAGACTTTTGTANAATAATTAATTATCATATCTTCCATTCTGTTATTGGGGGAGAAAT
21 >2559870H1 ADRETUT01 INCYTE
22 CACGAGGTCCTTCAGTTGAGACCAAGACCGGTGTGAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCA
23 TGAGGAGTGTGACTGTGTGTGTCAGAGGGAGCACAGGGGATAGCCGCATCACCACCAGCAGCTCTTGGCCAGAGCTGTGC
24 AGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGTCTCAAGGACCTTTCA
25 TCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGA
26 >3817470H1 BONSTUT01 INCYTE
27 TTAATAAGCAACTATGTTGCTATGAATTAACCTTGTGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAA
28 AATTTCTGCCATTTAGAGAAGAGAACTACATTCATGTTTGGAAAGAGATAAACCTGAAAAGAGAGTGGCCTTATCTTC
29 ACTTTATCGATAAGTCAGTTTATTTGTTTCATTTGTGTACATTTTATATTTCTCTTTTGACATTATAACTGTTGGCTTTC
30 TAATCTGTTAAATATATCTATTTTACCAAAGGTATTTAATATTTCTT
31 >3979767H1 LUNGTUT08 INCYTE
32 GGAGGATAGCCGCATCACCACCAGCAGCTCTTGGCCAGAGCTGTGTCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGC
33 GTTATCTCCATCCTTAATCTCAGTTGTTTGTCTCAAGGACCTTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAG
34 ACATCAAACAGAAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCATAAGGACAGGAGAAAGGTCTTCAATCGTG
35 GAAAGAAATTAATGTTGTATTAAATAGACACCAGCT
36 >3980011H1 LUNGTUT08 INCYTE
37 GGAGGATAGCCGCATCACCACCAGCAGCTCTTGGCCAGAGCTGTGTCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGC
38 GTTATCTCCATCCTTAATCTCAGTTGTTTGTCTCAAGGACCTTTTCATCTTCAGGATTTACATGCATTCTGAAAGAGGAGA
39 CATCAAACAGAAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCATAAGGACAGGAGAAAGGTCTTCAATCGTG
40 AAAGAAAATTAATGTTGTATTAAATAGATCACCA
41 >4925396H1 BLADDIT01 INCYTE
42 GAGAACCGATACCATTTTCTGGCCAGGTTGTCTCTGTTAAACGCTGTGGTGGGAACCTGTGCCTGTTGTCTCCACAATT
43 GCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAGTTGACACCAAAGACCGGTGTC
44 AGGGGATTGCAACAAATCACTCACCGACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGTCAGAGGGAGCACAGG
45 AGGATAGCCGCATCACCA
46 >3073703H1 BONEUNT01 INCYTE
47 AGAAAATCCAGAGTGGTGGATCTGAACCTTTTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCTCAGT
48 GTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAACCT
49 GTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCTCCCAAGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAG
50 TTGAGACCAAGACCGGTGTGAGGGGATTGCACAAATCA
51 >962169H1 BRAITUT03 INCYTE
52 AGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTTCATCTCGTATGGTGTAGAGTTAGATTAATCTGCA
53 TTTTAAAAAATGAAATTGGAATAGAATTGGTAAGTTGCAAGACTTTTGAATAAATTAATTAATATCATATCTTCCATTCT
54 CTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTATTCTTTT
55 GGGGAAATCTGAGCCTAGC
56 >4201385H1 BRAITUT29 INCYTE
57 TTTTAAAAAAGGAATATGTTGCTATGAATTAACCTTGTGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTAT
58 TAAAATTTCTGCCATTTAGAGAAGAGAACTACATTCATGTTTGGAAAGAGATAAACCTGAAAAGAAGAGTGGCCTATCT
59 TCACTTTATCGATAAGTCAGTTTATTTGTTTCATTTGTGTACATTTTATATTTCTCTTTTGACATATAACTGTTGGCTTT
60 CTAATCTGTTAAATATATCTATTTTACCAGGATTTAATAT
61 >1302516H1 PLACNOT02 INCYTE
62 AGGAAATCAATTAGGATAAGATTGTTGCTGATGAATATTTTCTCTGTAACCTTTCAACAGAGGAGGTAAGATTATAC
63 AGCTGCACACCTCGTAACCTTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCT
64 CCTGGTTAAACGCTGTGGTGGGAACGTGCTGCTCTCCACAATTGCAATGAATGTCAATGTGTCCTCAAGCAAAGTT
65 ACTAAAAAATACCACGAGGTC
66 >3684109H1 HEANOT01 INCYTE
67 ATTTTCATCTTCAGGATTTACAGTGCATTCTGAAANAGGAGCAAAATCAAACANAATTAGGAGTTGTGCAACAGCTCTTTTGA
68 GAGGAGGCCATAAGGACAGGAGAAAAGGCTCTTCAATCGTGGAAANAAAAATTAATGTTGTATTAAATAGATCACCAGCTA
69 GTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTTCAGTTCTTTGATACGGCTTAGGGTAATGTACAG
70 TACAGGAAAAAACTGTGCAAGTGAACACCTGATTCCGTTGCCTTGCTT
71 >2549720H1 LUNGTUT06 INCYTE
72 TTAGCTTGGNAATTTTCTAAACACAATGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAA
73 AATACATGATTTTCTCTCGTATGGTGTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGT
74 AAGTTGCAAGACTTTTGAATAAATTAATTATCATATCTTCCATTCTGTTATTGAGATGAAAAATAAAAAGCAACT
75 TATGANAGTAG

File: kony-est.mfas

76 >877279H1 LUNGAST01 INCYTE
 77 CTTTATGACAACCTTAGAACAACCTATTTTAGCTTGGTAAATTTTCTAAACACAATTGTTATAGCCAGAGGAACAAA
 78 GATG TAAAAATATTGTTGCTCTGACAAAAATACATGTAATTCATCTCGTATGGTGGCTAGAGTTAGATTAAATCTGCAT
 79 TTTAAAAAAGCTGAATTEGGAATAGAATTGGTAAGTTGCAAGGCTTTTGAATAATTAATTATCATATCTCCATTCC
 80 TGTTATTGGNGG
 81 >4713189H1 BRAINCT01 INCYTE
 82 CAAAGTTACTAAAAAATACCACGAGCTCTTTCAGTTGAGACCAAGACCGGTGTCACGGGATTGCACAAATCACTCACCG
 83 ACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTCAGAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAG
 84 CTCTTGCCCAGAGCTGTGCGAGTGCAGTGCTGATTCTATTAGAGACGATGCGTTATCTCCATCCTTAATCTCAGTTGT
 85 TTGCT
 86 >2171082H1 ENDCNOT03 INCYTE
 87 AGATAAACCTGAAAAGAGTGGCCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTA
 88 TATTCTCCTTTTGACATTATAACTGTGGCTTTTCTAATCTTGTTAAATATATCTATTTTACCAGGATTTTAATATT
 89 CTTTATGACAACCTTAGATCAACTATTTTAGCTTGGTAAATTTTCTAAACACAATTGTTATAGCCAGAGGAACAAA
 90 GATGA
 91 >875860H1 LUNGAST01 INCYTE
 92 CTGGATTTTTCATATTTCTTATTAAAAATTTCTGCCATTTAGAGAAGAGAACTACATTTCATGGTTTGGAGAGATAAAAC
 93 TGAAAAGAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTGTTTCATTGTGTACATTTTATATTCTCCT
 94 TTTGACATTATAACTGTTGGCTTTCTAATCTTGTTAAATATATCTATTTTACCAGGATTTTAATATCTTTTCTAT
 95 GAC
 96 >706168H1 SYNORAT04 INCYTE
 97 GCTCATATTCACATATGTAAACCAGAACATTTCTATGTACTACAAACCTGGTTTTTAAAAAGGANCTATGTTGCTATGAAT
 98 TAAACTTGTCTCGTGTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAAATTTCTGCCATTTAGAAGAAGAGAAC
 99 TACATTCATGGTTTGGAGAGATAAACCTGAAAAGAGAGTGGCCTTATCTTCANTTTATCGATAAGTCAGTTTATTTGT
 100 TTCA
 101 >458923H1 KERANOT01 INCYTE
 102 ANGAGTTGCCAGAGCTGTGTCAGTGCAGTGGCTGATTCTATTAGAGAACGATGCGTTATCTCCATCCTTAATCTCAGTT
 103 GTTGTNTCAAGGACCTTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAACAGAAATAGGAGTTGTG
 104 CAACAGCTCTTTTGAGAGGAGGCTAAAGGNCAGGAGAAAAGGCTTCAATCGTGGAAAGAAAATTAATGTTGTATTAA
 105 ATAGATC
 106 >538436H1 LNCDNOT02 INCYTE
 107 AAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCATTCTCGTATGGTGTAGAGTTAGATTAATCTG
 108 CATTTTAAAAAAGTGAATTTGGAATAGAATTGGTAAGTTGCAAGACTTTTTGAAAATTAATTAATTTATCATATCTTCCAT
 109 TCCTGTTATTGGAGATGAAAATAAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTAT
 110 >1303909H1 PLACNOT02 INCYTE
 111 AGGAAATCAAATTAGGATAAGATTGTATCTGATGAATATTTCTCTCTGAACCTTCTAACAGAGGAGGTAAGATTATAC
 112 AGCTGCACACCTCGTAACCTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCT
 113 CCTGGTTAAACGCTGTGGTGGGAAGTGTGCTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAG
 114 >2739211H1 OVARNOT09 INCYTE
 115 GTGCATTCTGAAAGAGGAGACATCAACAGAAATAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCCTAAAGGACAGGA
 116 GAAAAGGTCTTCAATCGTGGAAAGAAAATTAATGTTGTATTAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACG
 117 TATTCACATAGCTGGGTTCTGTATTTCAGTTCTTTGATACGGCTTAGGGTAATGTACAGGAAAAAACTGTGCAA
 118 GTGACACCTGAT
 119 >2550343H1 LUNGTUT06 INCYTE
 120 TGTACATTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCAATCTTGTTAAATATATCTATTTTACCAGG
 121 GTATTTAATATTCTTTTATGACAACCTTAGATCAACTATTTTAGCTTGGTAAATTTTCTAAACACAATTGTTATAGC
 122 CAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCATTCTCGTATGGTGCTA
 123 >5321148H1 FIBPFEN06 INCYTE
 124 CACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGNCAAAAAATACATGTATTTCATTCTCGTA
 125 TGGTGCTAGAGTTAGATTAAATCTGCATTTTAAAPAACTGAATTGGAATAGAATTGGTAAGTTGCAAGACTTTTGAAAA
 126 TAATTAATATATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAAGCAACTTATGAAAGTAAATTCAGATCCAC
 127 CATTACTAAC
 128 >879495H1 THYRNOT02 INCYTE
 129 ATTTCACTCTCGTATGGTGTAGAGTTAGATTAAATCTGCATTTTAAAAAAGTGAATTGGAATAGAATTGGTAAGTTGCAA
 130 AGACTTTTGAATAATTAATATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAAGCAACTTATGAAAGT
 131 AGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGGAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCT
 132 TGA AAAA
 133 >3325591H1 PTHYNOT03 INCYTE
 134 TGCAACAGCTCTTTTGAGAGGAGGCTTAAAGGACAGGAGAAAAGGCTTCAATCGTGGAAAGAAAATTAATGTTGTATT
 135 AAATAGATCACCAGCTAGTTTTCAGAGTTACCATGTACGATTTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACG
 136 GCTTAGGCTAATGTACAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCTTAAACCTTAAAGCNC
 137 ATGTCNNGGCNAANCGAAAAAT
 138 >543890H1 OVARNOT02 INCYTE
 139 TTTCTAAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTC
 140 TTCTCGTATGGTGCTAGAGTTAGATTAAATCTGCATTTTAAAAAAGTGAATTGGAATAGAATTGGTAAGTTGCAAGNCTT
 141 TTTGAAAATAAATAATATATATCTTCCATTCCTGTTATTGGAGGATGGAATAAATAAGCAACTTATGGAAAGTAGG
 142 ACATTCAGATC
 143 >3733565H1 SMCCNOS01 INCYTE
 144 CCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTTCATCTTACAGTGCATTCTGNAAGANGAGACATCAAAACAG
 145 AATTAGGNTTGTGCAAAAGCTCTTTTGAGAGGAGGCCATAAGGACAGGAGAAAAGGCTNCAATCGTGGAAAGNAAAT
 146 AAATGTTGTATNAAATNGATCACCAGCTAGTTTTCAGAGTTACCATGTACGATTTCCACTAGCTGGGNCNGTATTCACTCT
 147 TTCGGAACGCTTAGGGTAATGTACGTACAGGANAAAACTGTGCAAGTGTG
 148 >4641939H1 PROSTMT03 INCYTE
 149 GTACTACAAACCTGTTTTTAAAAAGGAACATGTTGCTATGAATTAACCTGTGTCCATGCTGATAGGACAGACTGGAT
 150 TTTNCATATTCTTATTAAAAATTTCTGCCATTTAGAAGAAGAGAACTACATTTCATGGTTTGGNAGAGATAAACCTGAAAA

[illegible]

226 CTTTTAAAAAAGTGAATGANTTGGTAAGTTGCAAAGNCNTTTGAAAATNAGTTATCAGAT
 227 >3530274H1 BLADNOT09 INCYTE
 228 TTCCCTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTATT
 229 CCTTCTGGGGAAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCG
 230 TGCTGTGCTGTGCAGTAGGAACACATCCTATTTATTGTGATGTGTGGTTTTATTATCTAAACTCTGTTCCATACACTTG
 231 TATAAATACATGGATATTTTATGTACAGAAGTATGTCTCTTAACCAAGTTCA
 232 >3530249H1 BLADNOT09 INCYTE
 233 CTTCATTCTCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGANAGTAGACATTCAGATCCAGCCATTACTAACCTAT
 234 TCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCG
 235 GTGCTGTGCTGTGCAGTAGGAACACATCCTATTTATTGTGATGTGTGGTTTTATTATCTTAARCTCTGTTCCATACACT
 236 TGIATAAATACATGGATATTTTATGTACAGAAGTATGTCTCTTAACCAAGTTCACTTATTGTACCTGG
 237

Fig 6 (cont'd)

VEGFE1	AAAATGTATGGATACAACTTAC	22
VEGFE2	GTTTGATGAAAGATTTGGGCTTG	23
VEGFE3	TTTCTAAAGGAAATCAAATTAG	22
VEGFE4	GATAAGATTTGTATCTGATG	20
VEGFE5	GATGTCTCCTCTTTCAG	17
VEGFE6	GCACAACTCCTAATTCTG	18
VEGFE7	AGCACCTGATTCCGTTGC	19
VEGFE8	TAGTACATAGAATGTTCTGG	20
VEGFE9	AAGAGACATACTTCTGTAC	19
VEGFE10	CCAGGTACAATAAGTGAACTG	21

Fig. 7

